Ann. Inst. Statist. Math.
32 (1980), Part A, 377-385

THE MINIMUM PROBABILITY ON AN INTERVAL WHEN THE MEAN AND VARIANCE ARE KNOWN

MORRIS SKIBINSKY

(Received June 6, 1978; revised June 18, 1980)

Abstract

This paper studies the minimum probability that distributions on a closed, bounded, non-degenerate interval can assign to its open subintervals when both the mean and variance are specified. It extends to this case Selberg's generalization of Tchebycheff's inequality.

1. Selberg's theorem

Let

$$J=[\alpha, \beta]$$

be a closed, bounded, non-degenerate interval. Denote by $V_J(\mu, \sigma^2)$ the class of all probability measures on J having mean μ and variance σ^2. It is well known that

$$V_J(\mu, \sigma^2) \text{ is non-empty} \Longleftrightarrow 0 \leqq \sigma^2 \leqq m_{\alpha,\beta}(\mu) ,$$

where for any numbers a, b, x we write

$$m_{a,b}(x)=(x-a)(b-x) .$$

Moreover

$$\sigma^2=0 \text{ or } m_{\alpha,\beta}(\mu) \Longleftrightarrow V_J(\mu, \sigma^2) \text{ is singleton} .$$

In [5] we studied in detail

$$\mathcal{U}_{a,b}^{(J)}(\mu, \sigma^2)=\max \{\mathrm{P}([a, b]):\ \mathrm{P} \in V_J(\mu, \sigma^2)\} \tag{1.1}$$

for all non-empty $V_J(\mu, \sigma^2)$ and each closed subinterval $[a, b]$ of J. Here we shall consider the minimum probability on open subintervals (a, b) of J which these same measures may achieve. (The minimum on closed subintervals is in general not attained). Thus for each open subinterval (a, b) of J, equivalently, for each pair of numbers a, b such that

(1.2) $$\alpha \leqq a < b \leqq \beta ,$$

define

(1.3) $$L_{a,b}^{(J)}(\mu, \sigma^2) = \min \{P((a, b)): P \in V_J(\mu, \sigma^2)\} .$$

For example, we have trivially that

$$L_{a,b}^{(J)}(\mu, 0) = I_{(a,b)}(\mu) , \qquad \alpha \leqq \mu \leqq \beta ,$$

where I denotes indicator function. The following theorem which shows explicitly the form that the minimum probability function (1.3) must assume was first proved by Selberg [4] with the real line $(-\infty, \infty)$ in place of J. As is evident below, this change makes no difference in the functional forms of the minimum probability since these forms do not depend on the endpoints of J. Selberg's results were rederived by Karlin and Studden in [2], pp. 475–479 and by Kemperman [3], p. 121, to illustrate how their general methodologies for obtaining such extrema might be applied. Isii [1], Theorems 1′ and 2′ (1 and 2) gives sharp lower (upper) bounds for the probability of open (closed) subsets of a bounded interval when the moments for an arbitrary Tchebycheff system of order m on that interval are specified. These result in a general methodology for deriving distributions of finite support which attain these bounds. Particularized to $m=2$, the classical power function Tchebycheff system, and interval subsets, these will also yield the theorem under consideration. In this connection, see also Theorem 2.1 and discussion p. 472 in [2]. We exhibit the theorem below with arbitrary J as here specified. A trivial modification of the approach in either [2] or [3] will suffice for proof which we delete. We use the following notation. For $a \leqq c \leqq b$ and any x let

(1.4) $$m_{a,c,b}(x) = \max [m_{a,c}(x), m_{c,b}(x)]$$

and let

(1.5) $$\rho = \rho_\mu(a, b) = a \text{ or } b \text{ according as } \mu < \text{ or } > (a+b)/2 .$$

THEOREM 1A. *For each open subinterval (a, b) of J and $\sigma^2 > 0$,*

$$L_{a,b}^{(J)}(\mu, \sigma^2) = \begin{cases} (\mu-\rho)^2/[\sigma^2 + (\mu-\rho)^2] , & \sigma^2 \leqq m_{a,(a+b)/2,b}(\mu) \\ 4[m_{a,b}(\mu) - \sigma^2]/(b-a)^2 , & m_{a,(a+b)/2,b}(\mu) < \sigma^2 \leqq m_{a,b}(\mu) \\ 0 , & m_{a,b}(\mu) < \sigma^2 \leqq m_{\alpha,\beta}(\mu) . \end{cases}$$

The conditions under which the three expressions for the minimum probability hold may also be viewed as restrictions on the endpoints of the open subinterval (a, b) for arbitrary μ, σ^2 such that

$$0<\sigma^2\leqq m_{\alpha,\beta}(\mu) . \tag{1.6}$$

As in [5], let

$$\tau(x)=\tau_{\mu,\sigma^2}(x)=\mu+\frac{\sigma^2}{\mu-x} , \qquad x\neq\mu . \tag{1.7}$$

See Lemmas 2.1, 2.2 and Figure 2.1 of [5] for simple properties of this function. Theorem 1A may be restated with conditions given in terms of τ.

THEOREM 1B. *Let μ, σ^2 satisfy* (1.6), *then for each open subinterval (a, b) of J*

$$L_{a,b}^{(J)}(\mu,\sigma^2)=\begin{cases}(\mu-\rho)^2/[\sigma^2+(\mu-\rho)^2] , & a\leqq\tau((a+b)/2)\leqq b\\ 4[m_{a,b}(\mu)-\sigma^2]/(b-a)^2 , & \tau(b)<(a+b)/2<\tau(a)\leqq b\\ 0 , & \textit{otherwise} .\end{cases}$$

As a conceptual convenience let us identify the open subintervals (a, b) of J one-one with the points (a, b) of the partly open triangle

$$T^0=\{(a, b):\ \alpha\leqq a<b\leqq\beta\} .$$

The conditions of Theorem 1B may then be viewed as a partition of T^0 into 3 sets parametrized by μ, σ^2 which satisfy (1.6).

The equivalence of respective conditions under Theorems 1A and B is easily established. Let

$$R=R(\mu,\sigma^2)=\{(a, b)\in T^0:\ a\leqq\tau((a+b)/2)\leqq b\} ,$$
$$S=S(\mu,\sigma^2)=\{(a, b)\in T^0:\ \tau(b)<(a+b)/2<\tau(a)\leqq b\} .$$

Intersecting R with sets in the T^0 partition whose members, $T_i=T_i(\mu)$, $i=1, 2,\cdots, 5$, are respectively determined by the 5 relationships:

$$\mu\leqq a ;\quad a<\mu<(a+b)/2 ;\quad (a+b)/2=\mu ;\quad (a+b)/2<\mu<b ;\quad \mu\leqq b ,$$

one finds

$$R\cap(T_1\cup T_3\cup T_5)=\phi ,$$

whereas

$$R\cap T_2=\{(a, b)\in T^0:\ a\leqq\tau((a+b)/2)<\mu\}=R_1 , \qquad \text{say} ,$$
$$R\cap T_4=\{(a, b)\in T^0:\ \mu<\tau((a+b)/2)\leqq b\}=R_2 , \qquad \text{say} ,$$

so that

$$R=R_1\cup R_2 .$$

Note that since $R_1\subset T_2$, $R_2\subset T_4$, the function ρ defined by (1.5) which appears in the first expression for the minimum probability is equal to a or b according as (a, b) is in R_1 or R_2. By (1.4)

$$0<\sigma^2\leqq m_{a,(a+b)/2,b}(\mu) \Longleftrightarrow 0<\sigma^2\leqq m_{a,(a+b)/2}(\mu) \text{ or } 0<\sigma^2<m_{(a+b)/2,b}(\mu) .$$

The first condition on the right-hand side is equivalent to $(a, b) \in R_1$, the second, to $(a, b) \in R_2$. Hence the condition on the left-hand side is equivalent to $(a, b) \in R$. Similarly,

$$m_{a,(a+b)/2,b}(\mu)<\sigma^2\leqq m_{a,b}(\mu) \Longleftrightarrow (a, b) \in S .$$

Finally, observe that

$$R\cup S=\{(a, b) \in T^0:\ m_{a,b}(\mu)\geqq\sigma^2\}=\{(a, b) \in T^0:\ a\leqq\tau(\beta),\ b\geqq\tau(a)\} .$$

A translation of Table 1 p. 475 in [2] to the notation and conditions here employed and simple manipulation yields

THEOREM 2. *Let* μ, σ^2 *satisfy* (1.6). *Let* ρ *and* τ *be respectively defined as in* (1.5) *and* (1.7) *and take* $c=(a+b)/2$, *then for each open subinterval*

$(a, b)\subset J$ such that	A distribution on J which has mean μ and variance σ^2 and assigns minimum probability to (a, b) is given by			
$a\leqq\tau(c)\leqq b$	support	ρ	$\tau(\rho)$	
	probability	$\frac{\tau(\rho)-\mu}{\tau(\rho)-\rho}$	$\frac{\mu-\rho}{\tau(\rho)-\rho}$	
$\tau(b)<c<\tau(a)<b$	support	a	c	b
	probability	$\frac{2(\mu-c)(\tau(c)-b)}{(b-a)^2}$	$1-\frac{4(\mu-c)(\tau(c)-c)}{(b-a)^2}$	$\frac{2(\mu-c)(\tau(c)-a)}{(b-a)^2}$

When neither condition holds (i.e. when the minimum probability is zero), it will suffice to take the distribution specified for the first condition replacing ρ *by* $\hat{\rho}$, *where*

$$\hat{\rho}=a \text{ or } \tau(\beta) \text{ according as } a< \text{ or } \geqq\tau(\beta) .$$

For $(a, b)\subset J$ such that $c=\mu$, which can occur under the second condition of the table, we interpret $(\mu-c)\tau(c)$ to be equal to its limit as $c\to\mu$, which is σ^2. Thus when $c=\mu$, the probabilities at a and b in the last row of the table are each equal to $2\sigma^2/(b-a)^2$; the probability at c, to $1-4\sigma^2/(b-a)^2$.

Define

$$M(\mu)=m_{\alpha,(\alpha+\beta)/2}(\mu) \text{ or } (\beta-\mu)^2/8 \qquad \text{according as } \mu\leqq \text{ or } \geqq(2\alpha+\beta)/3 ,$$

and let

$$\mathcal{M}(\mu)=M(\mu) \text{ or } M(\alpha+\beta-\mu) \qquad \text{according as } \mu\leqq \text{ or } \geqq(\alpha+\beta)/2 .$$

THEOREM 3. *Let* μ, σ^2 *satisfy* (1.6). *Then*

$$R_1(\mu, \sigma^2)=\phi \Longleftrightarrow \sigma^2>M(\mu) , \qquad R_2(\mu, \sigma^2)=\phi \Longleftrightarrow \sigma^2>M(\alpha+\beta-\mu) .$$

An immediate consequence of Theorem 3 and Theorem 1B is the

COROLLARY. *Let* μ, σ^2 *be such that*

$$\mathcal{M}(\mu)<\sigma^2\leqq m_{\alpha,\beta}(\mu) , \tag{1.8}$$

then for every open subinterval (a, b) *of* J

$$L_{a,b}^{(J)}(\mu, \sigma^2)=4[m_{a,b}(\mu)-\sigma^2]^+/(b-a)^2 ,$$

where f^+ *denotes* $f\cdot I_{\{f>0\}}$.

It should be noted that condition (1.8) becomes vacuous in the limit as the length of the interval J increases without bound, i.e., as $\beta-\alpha\to\infty$.

PROOF OF THEOREM 3. Since

$$R_1(\mu, \sigma^2)=\{(a, b)\in T^0:\ \sigma^2<m_{a,(a+b)/2}(\mu)\} ,$$

it suffices to verify that

$$\max_{(a,b)\in T^0} m_{a,(a+b)/2}(\mu)=M(\mu) .$$

The second equivalence follows by symmetry.

Examples. Let

$$J=[0, 4]$$

and suppose

$$\mu=3/2 , \qquad \sigma^2=25/16 .$$

We find

$$\mathcal{M}(3/2)=25/32<25/16\leqq 15/4=m_{0,4}(3/2) .$$

Hence by the corollary to Theorem 3,

$$L_{a,b}^{[0,4]}(3/2, 25/16)=4[m_{a,b}(3/2)-25/16]^+/(b-a)^2$$

for every open subinterval (a, b) of $[0, 4]$. Thus, for example, the minimum probability assignable to the interval $(1/2, 13/4)$ by any distribution on $[0, 4]$ with mean $3/2$ and variance $25/16$ is

$$4[(7/4)-(25/16)]/[(13/4)-(1/2)]^2=12/121 .$$

Now

$$[(1/2)+(13/4)]/2=15/8 , \qquad \tau(1/2)=49/16 ,$$

$$\tau(15/8)=-8/3 , \qquad \tau(13/4)=17/28 .$$

By Theorem 2, since

$$\tau(13/4)<15/8<\tau(1/2)<13/4 ,$$

a distribution on [0, 4] which has mean 3/2, variance 25/16, and assigns this minimum probability to the open interval (1/2, 13/4) is

$$\begin{matrix} 1/2 & 15/8 & 13/4 \\ 71/121 & 12/121 & 38/121 . \end{matrix}$$

Now suppose that

$$\mu=3/2 , \qquad \sigma^2=9/16 .$$

$R_2(3/2, 9/16)=\phi$ by Theorem 3 since $9/16>M(5/2)=9/32$, so that by Theorem 1B

$$L_{a,b}^{[0,4]}(3/2, 9/16)=\begin{cases} [(3/2)-a]^2/[(9/16)+((3/2)-a)^2] , & a\leqq\tau((a+b)/2)<3/2 \\ 4[m_{a,b}(3/2)-(9/16)]^+/(b-a)^2 , & \text{otherwise} . \end{cases}$$

For example to find the minimum probability assignable to the open interval (0.4, 3.8) by distributions on [0, 4] with mean 3/2 and variance 9/16, note that

$$(0.4+3.8)/2=2.1 \quad \text{and that} \quad 0.4<\tau(2.1)=9/16<3/2 ,$$

so that

$$L_{0.4,3.8}^{[0,4]}(3/2, 9/16)=(1.1)^2/[(9/16)+(1.1)^2]\cong 0.683 .$$

On the other hand, the minimum probability assignable to the open interval (0.4, 2.8), since

$$0.4\nless\tau(1.6)=-33/8 ,$$

is

$$4[(1.1)(1.3)-(9/16)]/(2.4)^2\cong 0.602 .$$

Finally, the minimum probability assignable to the open interval (0.4, 2) is zero. Distributions whose values at these intervals respectively attain these minimum probabilities are easily found via Theorem 2.

2. The minimum probability within k standard deviations of the mean

Let μ, σ^2 be arbitrary satisfying (1.6). We give here the minimum probability assignable to the open interval with endpoints $\mu \pm k\sigma$ ($k \geqq 0$, arbitrary) by any distribution on J possessed of this mean and variance. Equivalently, let X be a random variable distributed on J with mean μ and variance σ^2. We give the smallest value that

$$\mathrm{P}\,(|X-\mu|<k\sigma)$$

may have.

More precisely, let

$$L_k^* = L_k^*(\mu, \sigma^2) = \min\{\mathrm{P}\,((\mu-k\sigma, \mu+k\sigma)) : \mathrm{P} \in V_J(\mu, \sigma^2)\} .$$

THEOREM 4. *Let μ, σ^2 satisfy* (1.6). *Let*

$$A = \frac{1}{\sigma}\min(\mu-\alpha, \beta-\mu) , \qquad B = \frac{1}{\sigma}\max(\mu-\alpha, \beta-\mu)$$

and let

$$\Delta = A + B = (\beta-\alpha)/\sigma .$$

Then

$$L_k^* = \begin{cases} 0 , & 0 \leqq k \leqq \min(1, A) \\ (k^2-1)/k^2 , & 1 \leqq k \leqq A \\ (Bk+1)/\Delta\cdot(k+A) , & A < k \leqq 1/A \\ k^2/(k^2+1) , & \max(A, 1/A) < k \leqq B \\ 1 , & B < k . \end{cases}$$

Observe that either the second or the third condition must always be vacuous. Note that L_k^* is continuous on the left in k with jumps at $k=A$ and $k=B$. Note that as $A \to \infty$, only the standard Tchebycheff inequality remains.

PROOF. We shall suppose that

$$(\alpha+\beta)/2 \leqq \mu ,$$

equivalently that

$$A = (\beta-\mu)/\sigma , \qquad B = (\mu-\alpha)/\sigma .$$

The proof is strictly analogous when this inequality is reversed. For

P any probability measure on J,

$$\mathrm{P}\,((\mu-k\sigma,\ \mu+k\sigma))=\begin{cases}\mathrm{P}\,((\mu-k\sigma,\ \beta])\ , & A<k\leqq B\ ,\\ \mathrm{P}\,(J)=1\ , & B<k\ ,\end{cases}$$

so that

$$L_k^*=\begin{cases}L^{(J)}_{\mu-k\sigma,\,\mu+k\sigma}(\mu,\ \sigma^2)\ , & 0\leqq k\leqq A\\ 1-\mathcal{U}^{(J)}_{\alpha,\,\mu-k\sigma}(\mu,\ \sigma^2)\ , & A<k\leqq B\\ 1\ , & B<k\end{cases}\tag{2.1}$$

where $\mathcal{U}^{(J)}_{a,b}(\mu,\ \sigma^2)$ is defined by (1.1). Substituting $a=\mu-k\sigma$, $b=\mu+k\sigma$ for $k\leqq A$ into Theorem 1B, we find the condition under which the first expression for the minimum probability holds to be vacuous; the second expression, which reduces to $(k^2-1)/k^2$, to hold when $k\geqq 1$; zero to hold when $k<1$. Thus

$$L^{(J)}_{\mu-k\sigma,\,\mu+k\sigma}(\mu,\ \sigma^2)=0 \quad\text{or}\quad (k^2-1)/k^2$$

according as $0\leqq k\leqq \min(1,\ A)$ or $1\leqq k\leqq A$.

Corollary 1.2 of [5] gives $\mathcal{U}^{(J)}_{\alpha,b}(\mu,\ \sigma^2)$ for $\alpha\leqq b\leqq\beta$ and all μ, σ^2 satisfying (1.6). Substituting $b=\mu-k\sigma$ for $A<k<B$ into the conditions and expressions there given yields

$$\mathcal{U}^{(J)}_{\alpha,\,\mu-k\sigma}(\mu,\ \sigma^2)=\begin{cases}[A(k+\Delta)-1]/\Delta\cdot(k+A)\ , & A<k\leqq 1/A\\ 1/(k^2+1)\ , & \max(A,\ 1/A)<k\leqq B\ .\end{cases}$$

Subtraction from 1 yields the second line of (2.1) and the third and fourth lines of the theorem. The last line of the theorem is obvious. This completes the proof.

Examples. Let

$$J=[-3,\ 2]\ .$$

If

$$\mu=1/2\ ,\qquad \sigma^2=9/4\ ,$$

then $A=5/3$, $B=7/3$ and

$$L_k^*=\begin{cases}0\ , & 0\leqq k\leqq 1\\ (k^2-1)/k^2\ , & 1\leqq k\leqq 5/3\\ k^2/(k^2+1)\ , & 5/3<k<7/3\\ 1\ , & 7/3<k\ .\end{cases}$$

If

$$\mu=-2\ , \qquad \sigma^2=25/16\ ,$$

then $A=4/5$, $B=16/5$ and

$$L_k^*=\begin{cases} 0\ , & 0\leqq k\leqq 4/5 \\ (16k+5)/4(5k+4)\ , & 4/5<k\leqq 5/4 \\ k^2/(k^2+1)\ , & 5/4\leqq k\leqq 16/5 \\ 1\ , & 16/5<k\ . \end{cases}$$

Thus in particular there exist distributions on $[-3, 2]$ with mean $1/2$ and variance $9/4$ which place no mass at all within one standard deviation of their mean, but there exist *no* distributions with mean -2 and variance $25/16$ for which this is true. The least probability that any such distribution can place within one standard deviation of its mean is $7/12$.

The least probability that a distribution on $[-3, 2]$ with mean -2 and variance $25/16$ can place in any open interval with endpoints *more* than $4/5$ standard deviation from its mean is $89/160$, but there exists such a distribution which places no probability at all on the open interval with endpoints exactly $4/5$ standard deviation from its mean.

UNIVERSITY OF MASSACHUSETTS

REFERENCES

[1] Isii, K. (1960). The extrema of probability determined by generalized moments. I. Bounded random variables, *Ann. Inst. Statist. Math.*, **12**, 119-134.

[2] Karlin, S. and Studden, W. J. (1966). *Tchebycheff Systems; with Applications in Analysis and Statistics*, Interscience, New York.

[3] Kemperman, J. H. B. (1968). The general moment problem, a geometric approach, *Ann. Math. Statist.*, **39**, 93-122.

[4] Selberg, H. L. (1940). Zwei Ungleichungen zur Ergänzung des Techebycheffschen Lemmas, *Skand. Aktuarietidskrift*, **23**, 121-125.

[5] Skibinsky, M. (1977). The maximum probability on an interval when the mean and variance are known, *Sankhya*, **39**, A, 144-159.